

Gender diversity improves equity in peer review

Background

Peer review plays an important role in the assessment and communication of scientific research. In an ideal world, manuscripts are evaluated based on their content, with the traits of the scientist behind the work bearing no relevance. However, it has been suggested that implicit bias – the social and cultural stereotypes that can imperceptibly affect our daily judgements of others – may lead to peer review outcomes being influenced by, for example, the author's gender, nationality, or the country they work in (Lee et al., 2013. JASIST).

Founded in 2012 by the Howard Hughes Medical Institute, the Max Planck Society, and the Wellcome Trust, *eLife* is an open-access journal that publishes research in the life and biomedical sciences. Papers are first submitted as initial submissions and are assessed by a Senior Editor, usually in consultation with at least one member of the Board of Reviewing Editors (BRE). If a full submission is asked for, the BRE is encouraged to serve as one of the peer reviewers in addition to recruiting two to three external referees. When all the peer reviews have been submitted, the BRE and reviewers discuss the manuscript and their reports via an online forum. At this stage the identities of the BRE and the peer reviewers are known to one another. A more detailed description of the review process can be found in Schekman et al. (2013).

Building upon earlier research into eLife's review process (Giordan et al, 2016. F1000) and in light of concerns about implicit bias and gender diversity in scholarly communication (Helmer et al, 2017. eLife), we studied the effects of gender on decision outcomes at the journal.

Objectives

1. To contribute to ongoing research of implicit bias in scientific peer review;
2. To evaluate the effect of author and reviewer gender composition on encouragement and acceptance rates;
3. To determine whether reviewer gender diversity or homophily lead to fairer decisions more impartial outcomes.

Methods

We retrieved data on Research Articles, Short Reports, and Tools and Resources manuscripts submitted to *eLife* between 2012 and 2017; 23,732 initial submissions and 7,193 full submissions. For each initial submission we obtained the name, affiliation, and country of the corresponding author, and the name of the Senior Editor. In addition to this, we also obtained the names of the first author, last author, BRE, and the reviewers for manuscripts where there is a full submission.

We assigned the gender to each editor, reviewer, and author in our dataset, by taking the first name and country, and using an updated version of the algorithm used in Larivière et al. (2013). We then cross validated the gender assignment by applying it to dataset of names where the gender is known.

The encourage-acceptance rate is the percentage of full submissions that go on to be accepted.

To study gender homophily, we analyzed the gender composition of reviewer teams and the authors for full submissions. For each manuscript we designated the reviewer team gender as either all men, all women, mixed, uncertain; reviewer teams were labeled as mixed when we could identify the gender of at least one individual from each gender and as uncertain where it was uncertain if the reviewer composition falls into one of the above categories. On average, each full submission is reviewed by two to three reviewers. Also, the BRE may have served as one of these reviewers, and this is the case for 59.9% of full submissions. Of the 7,193 full submissions, 3,272 were reviewed by reviewer teams comprised only of men, 135 were reviewed by reviewer teams comprised of only women, 2,994 were reviewed by reviewers from both genders, and 792 where the reviewer composition was uncertain. More detailed methods are in Murray et al (in preparation).

Results

Authorship, gender, and outcomes.

From Figure 1, we can see that the majority of authors submitting are male, and where gender could be identified they represent 77% of last authors and 72.7% of first authors. We also observe a small, but significant, gender disparity that results in slightly more positive decision outcomes for submissions from male authors. Although not reported here, there is gender disparity in the percentage of initial submissions encouraged for review when considering the gender of the corresponding author – 30.5% (men) vs 28.6% (women), $\chi^2(1, n = 21,846) = 7.63, p < 0.01$. A similar effect is observed when comparing the percentage of initial submissions that are eventually accepted following review by the gender of the corresponding author – 15.4% (men) vs 13.6% (women), $\chi^2(1, n = 21,222) = 10.49, p < 0.01$. The greatest gender disparity we observe is in the relationship between the percentage of accepted full submissions (encourage-acceptance rate) and the gender of the last author $\chi^2(1, n = 6,036) = 5.66, p < 0.05$; where we see a disparity of 3.7 percent that leads to more positive decision outcomes for male authors. We see a similar gender disparity in the percentage of accepted full submissions by gender of the corresponding author, $\chi^2(1, n = 6,014) = 4.1, p < 0.05$. There are no significant gender disparities observed when considering the percentage of accepted full submissions and the gender of the first author, $\chi^2(1, n = 5,914) = 0.44, p > 0.05$. In summary, there is a small but statistically significant relationship between gender and submission outcomes for the corresponding and last author, with the most significant differences observed between the genders when taking into account the last author in full submissions. In each case the gender disparity favours men.

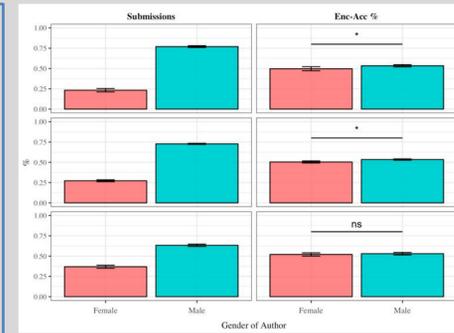


Figure 1 (above). Top row: Percentage of initial submissions and encourage-acceptance rate (Enc-Acc%) by gender of corresponding author; Middle row: Percentage of full submissions, and Enc-Acc% by gender of first author; Bottom row: Percentage of full submissions and Enc-Acc% by gender of last author. Those with unknown gender are excluded from analysis (Murray et al, in preparation). Vertical error bars indicate 95th confidence intervals of the proportion of submitted, encouraged, and accepted initial and full submissions. Asterisks indicate significance level of χ^2 tests of independence of frequency of encourage and acceptance by gender; **** = $p < 0.01$; *** = $p < 0.05$; "ns" = $p > 0.05$.

Gender homophily and peer review outcome

Figure 2 (below) groups the encourage-acceptance rates by gender of the author, for each authorship role, and by composition of the reviewer-team. For each group we examined the relationship between the gender of each authorship role and the outcome for the full submission. When a full submission is reviewed by an all-male or mixed-gender reviewer team, we observed that the rate of acceptance of full submissions is lower when the corresponding author is female. This effect however is smaller and not statistically significant in the case of mixed-gender review teams. When the reviewers are all-male, we see a statistically significant relationship between the gender of the last author and the acceptance rate of full submissions, which favours men, $\chi^2(1, n = 2,473) = 6.55, p < 0.05$; there is a similar effect when considering the gender of the corresponding author and the acceptance rate of full submissions, $\chi^2(1, n = 2,765) = 4.67, p < 0.05$. By comparison, we see no significant relationship between the gender of first author and acceptance rates when all reviewers are men; this also applies where we have all-female or mixed gender reviewer teams. For all all-female and mixed-gender reviewer teams, we see a trend that favours last, corresponding, and first female authors, albeit the numbers reported are low and not statistically significant. To summarise, for full submissions we see that male corresponding and last authors are more often accepted when reviewer by all-male reviewer teams, and that there is a greater gender parity in outcomes when reviewer teams are mixed.

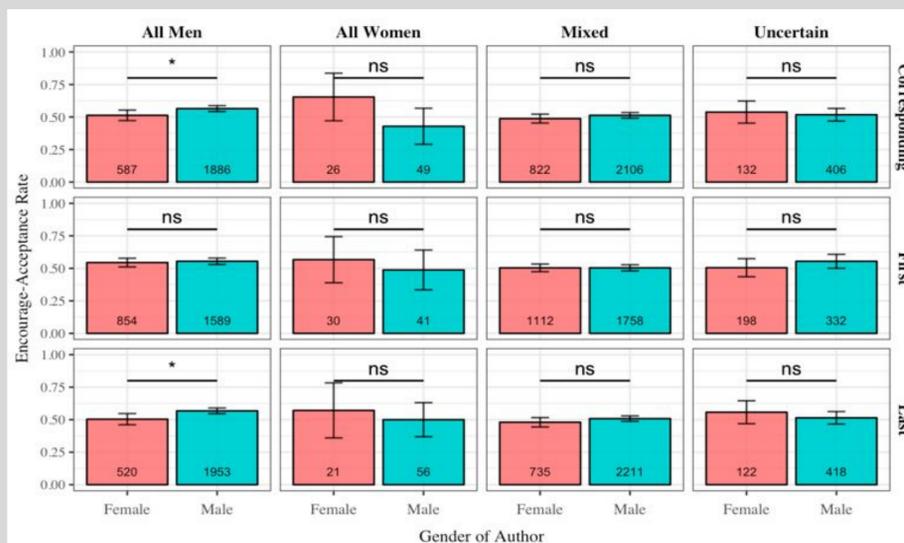


Figure 2 (left). Encourage-acceptance rates (acceptance rates of full submissions) by composition of team of reviewers and by gender of last, corresponding, and first authors (Murray et al, in preparation). Text at the base of each bar indicates the number of observations within each category of reviewer team and authorship gender. Vertical error bars indicate 95th percentile confidence intervals of the proportion of accepted full submissions. Asterisks indicate significance level of χ^2 tests of independence on frequency of acceptance by gender of author given each team composition; "ns" indicates no observed statistical significance. **** = $p < 0.01$; *** = $p < 0.05$.

Limitations

- The effects we observe may be specific to the eLife context. Given its consultative approach to peer review and the biological scope of the journal, our results may not be representative of peer review more generally.
- Because of the observational nature of the study, we cannot establish causal relationships between acceptance rates and author or gatekeeper demographics.
- The methodology used, especially the gender-algorithm, is a proxy and therefore not entirely accurate.

Conclusions

- Our results provide evidence that the gender of the corresponding author is associated with small yet significant differences in the rate at which initial submissions are encouraged (data not shown), and the rate at which full submissions are accepted at eLife. In both cases the disparity favours male authors.
- Our study also reveals the presence of a small yet significant differences in outcome based on the gender of the last author at eLife. Full submissions with male last authors are more likely to be accepted.
- We find that mixed-gender teams of reviewers at eLife produce more equitable outcomes between submissions with male and female lead authors. This finding supports the view that the gender of gatekeepers influences peer review outcomes.

References

- Giordan M, Csikasz-Nagy A, Collings AM, Vaggi F. (2016). The effects of an editor serving as one of the reviewers during the peer-review process. F1000Research 5:683. DOI: <https://f1000research.com/articles/5-683/v2>
- Helmer M, Schottdorf M, Neef A, Battaglia D. (2017). Research: Gender bias in scholarly peer review. eLife 6:e21718. DOI: <https://doi.org/10.7554/eLife.21718>
- Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. (2013). Bibliometrics: Global gender disparities in science. Nature News 504, 211-213. DOI: <https://doi.org/10.1038/504211a>.
- Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. Journal of the Association for Information Science and Technology, 64(1), 2-17. <https://doi.org/10.1002/asi.22784>
- Murray D, Siler K, Raymond J, Chan WM, Collings AM, Larivière V, Sugimoto CR. Gender and international diversity improves equity in peer review (in preparation)
- Schekman, R, Watt, F, Weigel D. 2013. Scientific Publishing: The eLife approach to peer review. eLife 2:e00799. DOI: <https://doi.org/10.7554/eLife.00799>

Acknowledgements: The results reported are part of a larger study on gender and international diversity, led by Cassidy Sugimoto, which is being prepared for submission. We thank James Gilbert for extracting and providing the raw data.

Funding: This work was funded in part by the Science of Science Innovation and Policy program of the National Science Foundation in the United States (NSF #1561299) to Cassidy Sugimoto.

Competing Interests: Wei Mun Chan, Susanna Richmond, Maria Guerreiro, and Andy Collings are employees of eLife.